CLAIMS

What is claimed is:

1. A method for generating a library of yeast expression vectors, comprising:

transforming into yeast cells

a linearized yeast expression vector having a 5'- and 3'-terminus sequence at the site of linearization, and

a library of insert nucleotide sequences that are linear and double-stranded, each insert sequence comprising

a first nucleotide sequence encoding a first polypeptide subunit,

a second nucleotide sequence encoding a second polypeptide subunit,

a linker sequence encoding a linker peptide that links the first and second polypeptide subunits, and

a 5'- and 3'- flanking sequence at the ends of the insert sequence which are sufficiently homologous to the 5'- and 3'-terminus sequences of the linearized yeast expression vector, respectively, to enable homologous recombination to occur; and

having homologous recombination occur between the vector and the insert sequence such that the insert sequence is included in the vector in the transformed yeast cells,

wherein

the first polypeptide subunit, the second polypeptide subunit, and the linker polypeptide are expressed as a single fusion protein; and

the first and second nucleotide sequences each independently varies within the library of expression vectors.

2. The method of claim 1, wherein the 5'- or 3'- flanking sequence of the insert nucleotide sequence is between about 30-120 bp in length.

3. The method of claim 1, wherein the 5'- or 3'- flanking sequence of the insert nucleotide sequence is between about 40-90 bp in length.

4. The method of claim 1, wherein the 5'- or 3'- flanking sequence of the insert nucleotide sequence is between about 60-80 bp in length.

5. The method of claim 1, wherein the linker sequence of the insert nucleotide sequence is between 30-120 bp in length.

6. The method of claim 1, wherein the linker sequence of the insert nucleotide sequence is between 45-102 bp in length.

7. The method of claim 1, wherein the linker sequence of the insert nucleotide sequence is between 45-63 bp in length.

8. The method of claim 1, wherein the linker sequences of the insert nucleotide sequence comprises a nucleotide sequence encoding an amino acid sequence of Gly-Gly-Gly-Gly-Ser in 3 or 4 tandem repeats.

9. The method of claim 1, wherein the yeast expression vector is a 2μ plasmid vector.

10. The method of claim 1, wherein the diversity of the first or the second polypeptide subunit within the library of fusion proteins is at least 10^3.

11. The method of claim 1, wherein the diversity of the first or the second

polypeptide subunit within the library of fusion proteins is at least 10^4.

12. The method of claim 1, wherein the diversity of the first or the second polypeptide subunit within the library of fusion proteins is at least 10^5.

13. The method of claim 1, wherein the diversity of the fusion proteins encoded by the library of yeast expression vectors is at least $1x10^6$.

14. The library of claim 1, wherein the diversity of the fusion proteins encoded by the library of yeast expression vectors is at least $1x10^{10}$.

15. The method of claim 1, wherein the diversity of the fusion proteins encoded by the library of yeast expression vectors is at least $1x10^{12}$.

16. The method of claim 1, wherein the diversities of the first and second polypeptide subunits are each independently derived from libraries of precursor sequences that are not specifically designed for a target peptide or protein.

17. The method of claim 1, wherein the diversities of the first and second polypeptide subunits are not derived from one or more proteins that are known to bind to a target peptide or protein.

18. The method of claim 1, wherein the diversities of the first and second polypeptide subunits are not generated by mutagenizing one or more proteins that are known to bind to a target peptide or protein.

19. The method of claim 1, wherein the first nucleotide sequence is 5' relative to the second nucleotide sequence.

20. The method of claim 19, wherein the first nucleotide sequence in the library of expression vectors comprises a coding sequence of an antibody heavy-chain variable region, and the second nucleotide sequence comprises a coding sequence of an antibody light-chain variable region.

21. The method of claim 1, wherein the linker peptides expressed by the library of expression vectors provide a substantially conserved conformation between the first and second polypeptide subunits across the library of fusion proteins expressed by the library of expression vectors.

22. The method of claim 1, wherein the conformation of the fusion protein having the first and second polypeptide subunits linked by the linker peptide mimics a conformation of a single chain antibody.

17

~~23.~~ The method of claim 1, wherein each of the expression vectors further comprises a sequence encoding an affinity tag.

24. The library of claim 23, wherein the affinity tag is selected from the group consisting of a polyhistidine tag, polyarginine tag, glutathione-S-transferase, maltose binding protein, staphylococcal protein A tag, and an EE-epitope tag.

25. A method for generating a library of yeast expression vectors, comprising:

a) transforming into yeast cells

i) a linearized yeast expression vector having a 5'- and 3'-terminus sequence at a first site of linearization; and

ii) a library of first insert nucleotide sequences that are linear,

double stranded, each of the first insert sequences comprising a first nucleotide sequence encoding a first polypeptide subunit, a 5'- and 3'-flanking sequence at the ends of the first insert sequence which are sufficiently homologous to the 5'- and 3'-terminus sequences of the vector at the first site of linearization, respectively, to enable homologous recombination to occur;

b) having homologous recombination occur between the vector and the first insert sequence in the transformed yeast cells, such that the first insert sequence is included in the vector;

c) isolating from the transformed yeast cells the vectors that contain the library of the first insert sequences;

d) linearizing the vectors containing the library of the first insert sequences to generate a 5'- and 3'- terminus sequence at a second site of linearization;

e) transforming into yeast cells

i) the linearized yeast expression vectors in step d), and

ii) a library of second insert nucleotide sequences that are linear, double stranded, each of the second insert sequences comprising a second nucleotide sequence encoding a second polypeptide subunit, a 5'- and 3'-flanking sequence at the ends of the second insert sequence which are sufficiently homologous to the 5'- and 3'-terminus sequences of the vector at the second site of linearization, respectively, to enable homologous recombination to occur; and

f) having homologous recombination occur between the linearized yeast expression vector at the second linearization site and the second insert sequences in the transformed yeast cells, such that the second insert sequence is included in the vector and the first and second nucleotide sequences are linked by a linker sequence;

wherein

the expression vector expresses the first polypeptide subunit, the second polypeptide subunit, and the linker polypeptide as a single fusion protein; and

the first and second nucleotide sequences each independently varies within the library of expression vectors.

26. The method of claim 25, wherein the 5'- or 3'- flanking sequence of the first or the second insert nucleotide sequence is between about 30-120 bp in length.

27. The method of claim 25, wherein the 5'- or 3'- flanking sequence of the first or the second insert nucleotide sequence is between about 40-90 bp in length.

28. The method of claim 25, wherein the 5'- or 3'- flanking sequence of the first or the second insert nucleotide sequence is between about 60-80 bp in length.

29. The method of claim 25, wherein the linker sequence is between 30-120 bp in length.

30. The method of claim 25, wherein the linker sequence is between 45-102 bp in length.

31. The method of claim 25, wherein the linker sequence is between 45-63 bp in length.

32. The method of claim 25, wherein the linker sequence comprises a nucleotide sequence encoding an amino acid sequence of Gly-Gly-Gly-Gly-

Ser in 3 or 4 tandem repeats.

33. The method of claim 25, wherein the 5'- and 3'-flanking sequences at the ends of the first or second insert nucleotide sequence comprise a 5'- and 3'-site-specific recombination site, respectively, that are recognized by a site-specific recombinase.

34. The method of claim 33, wherein one of the 5'- and 3'-site-specific recombination sites is coliphase P1 loxP, and the other is a mutant loxP sequence.

35. The method of claim 33, wherein the 5'- and 3'-site-specific recombination sites are each independently selected from the group consisting of SEQ ID Nos 1-13.

36. The method of claim 33, wherein the site-specific recombinase is CRE recombinase.

37. The method of claim 33, further comprising:

causing site-specific recombination between the members of the library of the yeast expression vectors at the 5'- and 3'-recombination sites, the recombination resulting in exchange of the first or second nucleotide sequences between the members of the library of the yeast expression vectors.

38. The method of claim 37, wherein the recombination is caused by expression of a recombinase that is inducibly controlled in the yeast cells.

39. The method of claim 38, wherein the 5'- and 3'-recombination sites are

different loxP sequences, and the recombination is caused by inducible expression of CRE recombinase in the yeast cells.

40. A method of producing a library of single chain antibodies, comprising:

expressing in yeast cells a library of yeast expression vectors, each vector comprising

a first nucleotide sequence encoding an antibody heavy chain variable region,

a second nucleotide sequence encoding an antibody light chain variable region, and

a linker sequence encoding a linker peptide that links the antibody heavy chain variable region and the antibody light chain variable region,

wherein

the antibody heavy chain variable region, the antibody light chain variable region, and the linker peptide are expressed as a single fusion protein; and

the first and second nucleotide sequences each independently varies within the library of expression vectors to generate a library of single-chain antibodies with a diversity of at least 10^6.

41. The method of claim 40, wherein the diversity of the library of single-chain antibodies is between about $1x10^6$-$1x10^{18}$.

42. The method of claim 40, wherein the diversity of the library of single-chain antibodies is between about $1x10^8$-$1x10^{18}$.

43. The library of claim 40, wherein the diversity of the library of single-chain antibodies is between about $1x10^{12}$-$1x10^{18}$.